



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07C 237/00, C07H 21/00, C07K 1/00, 5/00, 7/00, 13/00, 17/00	A1	(11) International Publication Number: WO 94/13623 (43) International Publication Date: 23 June 1994 (23.06.94)
(21) International Application Number: PCT/US93/12013 (22) International Filing Date: 10 December 1993 (10.12.93) (30) Priority Data: 07/988,278 11 December 1992 (11.12.92) US (71) Applicant: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608-2916 (US). (72) Inventors: ZUCKERMANN, Ronald, N.; 2534 Grant Street, Berkeley, CA 94703 (US). KERR, Janice, M.; 3863A Greenwood Avenue, Oakland, CA 94602 (US). (74) Agents: BOZICEVIC, Karl et al.; Morrison & Foerster, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).		(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, LV, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: SYNTHESIS OF ENCODED POLYMERS <div style="text-align: center;"> </div> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="width: 45%;"> <ul style="list-style-type: none"> • chemical functionality • "binding" strand • novel building blocks </div> <div style="width: 45%;"> <ul style="list-style-type: none"> • information storage • "coding" strand • DNA or peptide </div> </div>		
(57) Abstract <p>Conjugates and methods of producing the conjugates and mixtures thereof are disclosed. Conjugates are comprised of an active polymer made up of monomer units selected from the group consisting of the monomer units of peptides and/or peptoids, and an encoding polymer comprised of encoding monomers wherein the encoding polymer corresponds to the active polymer, and a coupling moiety covalently coupled to the active polymer and the encoding polymer. In accordance with the synthesis methodology, mixtures of large numbers of conjugates are produced by providing a coupling moiety and covalent binding it to an active monomer and an encoding monomer. Additional monomer units are added to the active monomer to create an active polymer and additional monomer units are added to the encoding monomer to produce an encoding polymer until the desired length is reached for the active polymer. Mixtures of conjugates attached to support bases can be used to assay a sample. The sample is brought into contact with the conjugates and a determination is made with respect to which active proteins bind to receptor sites within the sample. When active binding proteins are determined, the encoding polymer associated with the active polymer is sequenced, and by deduction, the sequence of the active polymer is determined.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

-1-

5

SYNTHESIS OF ENCODED POLYMERSField of the Invention

This invention relates to the fields of biopolymer synthesis and drug design. More particularly, the invention relates to methods for synthesizing libraries of biologically active polymers in association with an included polymer which is encoded to facilitate deciphering.

15 Background of the Invention

Modern pharmaceutical technology has taken two divergent paths in pursuit of new therapeutic compounds. Rational drug design achieves results by intensive analysis of the molecular structure of binding sites, and designing compounds specifically to complement a desired binding site. For example, one interested in preparing new antihypertensive compounds might analyze the molecular structure of the β -adrenergic receptor binding site using X-ray crystallography and/or advanced NMR techniques, and then synthesize compounds calculated to fit within the binding site and complement the charge distribution.

The other approach is to prepare an enormous library of compounds and select only those compounds which exhibit a desired activity. This approach differs from the traditional pharmaceutical cycle of design/synthesize/test/synthesize variants by conducting the screening step in a massively parallel fashion, screening an enormous number of different compounds simultaneously. The challenge to this

-2-

approach is first to provide a group of compounds for screening that is sufficiently numerous and diverse to insure that the activity sought is represented in the group, and second to identify the active compounds at low concentration within the group.

Rutter et al., US 5,010,175 disclosed a method of making diverse mixtures of peptides by adjusting the concentration of each activated peptide in proportion to its reaction rate, in order to obtain a substantially equimolar mixture of peptides. Rutter also disclosed the process of providing a mixture of peptides (having at least 50 different peptides), and selecting one or more peptides having a desired property and separating them from the rest of the peptides.

Zuckermann et al., PCT WO91/17823 disclosed an alternative method for preparing diverse mixtures of oligopeptides on solid-phase resins, and a robotic device for performing the necessary manipulations. In this method, a pool of resin particles is separated into a number of groups (wherein each group is defined as one or more separate reactions), and a different amino acid coupled to the resin in each group. The groups are then mixed together, separated into a number of groups, and again coupled with a different amino acid for each group. This cycle is repeated until the desired number of amino acids per oligopeptide is obtained. One advantage of this approach is that each coupling reaction occurs in isolation from other reactants, which permits one to drive each reaction to completion without carefully adjusting the initial concentration of each reactant. This method also facilitates the preparation of oligopeptides wherein some positions within the peptide chain are held constant, and where some positions are restricted to less than all amino acids. For example, one may use this method to prepare a peptide of the formula $X_1-X_2-X_3\text{-Glu-Ala-}X_4\text{-}X_5\text{-}X_6$, where X_n

-3-

can be any amino acid. If desired, one could limit, for example, X_3 and X_5 to hydrophobic residues.

Zuckermann also disclosed that this method may be applied to the synthesis of oligonucleotides, which
5 may then be inserted into cloning and expression vectors for biological expression.

Bartlett et al., PCT WO91/19735 disclosed a variation of the Zuckermann et al. method in which a diverse set of non-amino acid monomers is employed to
10 form mixtures of compounds called "peptoids." Peptoids sample a different region of physico-chemical parameter space than traditional oligopeptides, depending on the type of linkage between monomers, and may be able to exhibit activities unavailable to
15 peptide libraries due to the diversity (or difference) in side chains.

Houghten, US 4,631,211 disclosed a "tea-bag" peptide synthesis method. The "tea bags" are mesh bags containing resin beads for peptide synthesis.
20 Houghten's method enables one to add the same amino acid to a number of different oligopeptides without mixing the products: a number of "tea bags" may be reacted with an amino acid in a common pot, then separated physically.

25 Cook, EP 383620 described synthesis of COP-1, a random polymer of Ala, Glu, Lys, and Tyr, having an average molecular weight of 23 kDa having activity in the treatment of multiple sclerosis. COP-1 is made in the prior art by chemical polymerization of the
30 amino acids. However, Cook described expression from genes made by random polymerization of oligonucleotides, and selection for those clones expressing COP-1 with the highest activity.

Lebl et al., EP 445915 described a machine
35 for performing multiple simultaneous peptide syntheses using a planar support surface. The planar support is, for example, paper or cotton.

-4-

Kauffman et al., WO86/05803 disclosed production of peptide libraries by expression from synthetic genes which are partially or wholly "stochastic." Stochastic genes are prepared by polymerizing a mixture of at least three oligonucleotides (at least heptamers) to form a double-stranded stochastic sequence, and ligating the stochastic sequence into an expression vector.

Lam et al., WO92/00091 disclosed libraries of oligonucleotides, oligopeptides, and peptide/nucleotide chimeras, and methods for screening the libraries for active compounds. However, Lam did not disclose conjugates having an active sequence and a coding sequence.

K.M. Derbyshire et al., Gene (1986) 46:145-52 disclosed a method for "saturation mutagenesis" of a segment of DNA, by synthesizing oligonucleotides using contaminated pools of monomer. Each A, C, G, and T reservoir contained 1/54 parts of each of the other bases. The object was to prepare a DNA segment mixture having one or two mutations per sequence. They did not observe equal frequencies of mutation, presumably due to differences in coupling efficiency. The authors suggested synthesizing sequences using four reservoirs containing pure bases, and one reservoir containing a mixture of all four bases in the concentrations necessary to balance the coupling efficiencies.

J.F. Reidhaar-Olson et al., Science (1988) 241:53-57 disclosed the generation of mutant λ repressor proteins by replacing two codons with random nucleotides (NNG/c). The resulting mutant proteins were assayed for activity to determine which amino acid positions were critical, and which positions should be conserved.

I.S. Dunn et al., Prot Eng (1988) 2:283-91 disclosed the use of random polynucleotides to generate mutant β -lactamase α -peptides, some of which

-5-

exhibited properties superior to the native sequence α -peptide.

A.R. Oliphant & K. Struhl, Nuc Acids Res (1988) 16:7673-83 disclosed the use of random polynucleotides to investigate promoter function. A section of random polynucleotide was inserted into the -35 to -10 region of a gene conferring drug resistance in *E. coli*, and the transformants screened for resistance. Survivors were cloned and sequenced to provide a functional consensus sequence.

F.W. Studier, Proc Natl Acad Sci USA (1989) 86:6917-21 disclosed a method for sequencing large volumes of DNA by random priming of cosmid libraries.

A.R. Oliphant et al., Proc Natl Acad Sci USA (1989) 86:9094-98 disclosed the generation of β -lactamase mutants having altered properties, by cloning a random polynucleotide into the β -lactamase gene.

D.K. Dube et al., Biochem (1989) 28:5703-07 disclosed the generation of β -lactamase mutants having altered properties, by cloning a random polynucleotide into the β -lactamase gene.

R.A. Owens et al., Biochem Biophys Res Comm (1991) 181:402-08 disclosed the selection of an HIV protease inhibitor from a library of 240,000 tetrapeptides (in 22 mixtures). The mixtures were prepared by the "mixed resin" technique.

These techniques enable one to prepare libraries of diverse compounds. However, the problem of identifying the resulting compounds has seldom been addressed. Oligopeptides are typically sequenced by stepwise cleavage of each amino acid from the parent compound (which is usually immobilized on a resin), with chromatographic analysis of the cleaved moiety. Sensitive techniques are required to distinguish between twenty or more amino acids. Analysis is further complicated when uncommon amino acids are

-6-

employed (using current techniques), especially when monomers are linked without using amide bonds.

Summary of the Invention

5 The present invention provides a method of synthesizing true mixtures of diverse oligopeptides and/or peptide-like compounds along with an associated encoding polymer making it possible to easily analyze those compounds exhibiting a desired activity. The
10 invention involves synthesizing an encoding DNA strand simultaneously with the peptide/peptoid. Each unique peptide/peptoid sequence associated with its own unique DNA strand to provide the conjugates of the invention. These conjugates are screened to determine
15 which peptide/peptoid compounds exhibit a desired activity, and the active conjugates analyzed by DNA sequencing methods to determine the attached peptide/peptoid sequence by deduction, i.e., since each DNA sequence is
20 associated with a known peptide/peptoid, once the DNA sequence is determined, the sequence of the peptide/peptoid can be deduced.

 Another aspect of the invention is a
25 conjugate comprising a peptide or peptoid coupled to and/or directly associated with a coding polymer (CP), e.g. a nucleic acid (NA). The peptide/peptoid/CP conjugate may be linked directly (i.e., covalently bound either directly or through a small organic molecule), or by linkage to the same support (e.g., by
30 synthesizing both peptide/peptoid and CP strand on the same particle or bead of resin).

 An important object of the invention is to provide a chemical synthesis method which allows the
35 production of libraries of peptides and/or peptoids along with a unique encoded polymer such as a DNA strand which makes it possible to readily determine the sequence of the peptide or peptoid.

-7-

An advantage of the present invention is that the methodology makes it possible to readily identify and sequence peptides and/or peptoids having desirable biological activities.

5 A feature of the present invention is that sequences of peptoids or peptides which contain nonconventional amino acids can still be readily determined by sequencing associated polymers such as DNA sequences which are simultaneously synthesized
10 with the peptoids and encode them.

These and other objects, advantages and features of the present invention will become apparent to those persons skilled in the art upon reading the details of the structure, synthesis and use as more
15 fully set forth below, reference being made to the accompanying figures forming a part hereof.

Brief Description of the Drawings

Figure 1 is a schematic diagram showing a
20 specific embodiment of a conjugate of the invention which conjugate includes a "binding" strand or active polymer attached to a solid-support substrate which substrate is also bound to an information storage or "coding" strand;

25 Figure 2 is a schematic flow diagram demonstrating how encoded libraries can be synthesized on beads as the solid-support substrate;

Figure 3 is a schematic diagram showing methods of the synthesis of both solid-phase and
30 solution-phase libraries;

Figure 4 is a schematic diagram showing resin-bound libraries generated by the derivatization of non-hydrolyzable resins;

35 Figure 5 is an HPLC chromatogram of binding and coding peptide strands simultaneously synthesized via non-hydrolyzable resin linkage;

-8-

Figure 6 is an HPLC chromatogram of a coding and binding strand adduct which was synthesized via a hydrolyzable resin linker;

5 Figure 7 is a plotted graph resulting from ELISA competition of binding sequences versus binding/encoding sequences;

Figure 8 is a schematic diagram showing the analysis of a solid-phase amptide; and

10 Figure 9 is a schematic flow diagram showing the analysis of a solution-phase amptide.

Detailed Description of Preferred Embodiments

Before the present method of synthesis, conjugates and methods of using such are described, it
15 is to be understood that this invention is not limited to the particular methodologies, conjugates, or methods of use described as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular
20 embodiments only and is not intended to be limiting since the scope of the present invention will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "and,"
25 and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a peptide" includes mixtures of peptides, reference to "an amino acid" includes mixtures thereof, and reference to "the reaction"
30 includes one or more reactions of the same type as generally understood by those skilled in the art, and so forth.

Unless defined all otherwise, all technical and scientific terms used herein have the same meaning
35 as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or

-9-

testing of the present invention, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to describe and disclose specific information for which the reference was cited in connection with.

In general, the invention provides a rapid method of synthesizing large numbers of conjugates which conjugates are comprised of a peptide/peptoid sequence, e.g., an amino acid sequence associated with a unique encoding sequence, e.g., a DNA sequence. The conjugates can be readily synthesized and thereafter screened for biological activity, and when activity is found, the particular peptide/peptoid sequence found to be active can be readily identified by its associated encoding (DNA) strand. Each conjugate of the invention is comprised of at least two components with one of the components being the peptide or peptoid sequence which binds to a receptor of interest and the other sequence being a polymer which encodes the binding sequence. The invention may utilize standard amino acids and DNA as encoding monomers to produce a chemically diverse library of solution-phase or solid-phase conjugates. In order to further describe the invention in detail, the following definitions are provided.

A. Definitions

The terms "nucleic acid" and "NA" refer to oligomers constructed from DNA and/or RNA bases which may be sequenced using standard DNA sequencing techniques. The NAs used herein may include uncommon bases so long as such bases are distinguishable from the other bases employed under the DNA sequencing methods to be used and include peptide-nucleic acids (PNAs) (disclosed by Nielsen, P.E., Egholm, M., Berg, R.H. & Buchardt, O., Science (1991) 254, 1497-1500). Such PNAs could serve as coding strands and the

-10-

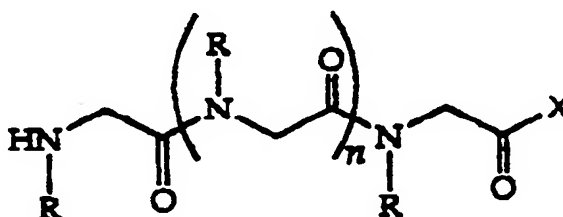
detection would be by hybridization. NAs will usually be constructed from monomers linked by phosphodiester bonds, but other similar linkages may be substituted if desired. For example,

5 phosphorothioates may be employed to reduce lability.

The term "peptide" as used herein refers to the 20 commonly-occurring amino acids: alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H),
 10 isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V), tryptophan (W), and tyrosine (Y).

The term "peptoid" as used herein refers to a non-peptide monomer of the general formula $(R)_n-X-(L)_m$, where R is a side chain group, n is at least 1, L is a linking group, m is at least 2, and X is a small organic radical. It is preferred to select L radicals that may be individually protected and
 20 deprotected. Preferably n will be 1 or 2 and m will be 2. Monomers wherein m is 3 or greater may be used to form branched active polymers. Presently preferred monomers are N-substituted glycine derivatives of the formula

25



30 wherein R is alkyl of 2-6 carbon atoms, haloalkyl of 1-6 carbon atoms wherein halo is F, Cl, Br, or I, alkenyl of 2-6 carbon atoms, alkynyl of 2-6 carbon atoms, cycloalkyl of 3-8 carbon atoms, alkoxyalkyl of 2-8 carbon atoms, aryl of 6-10 carbon atoms, arylalkyl
 35 of 7-12 carbon atoms, arylalkyl of 7-12 carbon atoms substituted with 1-3 radicals independently selected from halo and nitro and hydroxy, aminoalkyl of 1-6 carbon atoms, hydroxyalkyl of 1-6 carbon atoms,

-11-

carboxy, carboxyalkyl of 2-6 carbon atoms, carboalkoxy-alkyl of 3-10 carbon atoms, carbamyl, carbamylalkyl of 2-6 carbon atoms, imidazolyl, imidazolylalkyl of 4-10 carbon atoms, pyridyl, pyridylalkyl of 6-10 carbon atoms, piperidyl, piperidylalkyl of 5-10 carbon atoms, indolyl, or indolylalkyl of 9-15 carbon atoms. Thus, active polymers composed of these monomers are equivalent to polyglycine having side chains attached at each nitrogen. These and other monomers are described in copending application USSN 07/715,823, incorporated herein by reference, and PCT WO91/19735.

The terms "coding" and "encoding" indicated that one or more coding monomers corresponds directly and uniquely to a given active monomer, e.g., conventional nucleic acids encode (in groups of three) the 20 natural amino acids. The number of coding monomers used for each code depends on the number of different coding monomers and the number of different active monomers. Typically, the number of different active monomers used will range from about 5 to about 30. A basis set of 4 coding monomers can encode up to 16 active monomers taken in "codons" of 2 coding monomers. By increasing the coding monomer basis set to five distinct monomers, one can encode up to 25 different peptide/peptoid monomers. A basis set of 4 coding monomers can encode up to 64 peptide/peptoid monomers taken in "codons" of 3 coding monomers. Note that one can make the code degenerate or nondegenerate, and can insert additional coding information into the sequence. For example, one may wish to begin each codon with the same base (e.g., G), using that base only in the first position, thus unambiguously identifying the beginning of each codon. As a practical matter, the group of monomers selected for use as coding monomers will form polymers that are easier to

-12-

sequence than the active polymers, i.e., the coding monomers may be more readily identified using present day sequencing technology as compared to the monomer of the active polymers. With current technology, the order of preference for coding monomers is nucleic acids > peptides > peptoids. Nucleic acids have the additional advantage that the coding sequence may be amplified by cloning or PCR (polymerase chain reaction) methods known in the art.

10 The term "active polymer" and/or "binding polymer" refers to a polymer having a desired biological activity. Suitable biological activities include binding to natural receptors, pharmaceutical effects, immunogenicity/antigenicity, and the like.

15 "Immunogenicity" refers to the ability to stimulate an immune response (whole or partial serum-mediated immunity and/or cell-mediated immunity) in a bird or mammal following administration. Antigenicity requires only that the active polymer bind to the

20 antigen-binding site of an antibody. Pharmaceutical activities, for the purposes of this invention, will generally depend on the ability of the active polymer to bind a protein, carbohydrate, lipid, nucleic acid, or other compound present in the subject. For

25 example, an active polymer may bind to a cell surface receptor and compete with the receptor's natural ligand, with or without activation of the receptor. Other useful pharmaceutical activities include cleavage of endogenous molecules (e.g., protease

30 activity, nuclease activity, and the like), catalysis of reactions either primarily or as a cofactor, donation of functional groups (e.g., acyl, ATP, alkyl, and the like), pore formation, and the like. Active polymers comprise a series of monomers which are

35 linked sequentially. The monomers will generally be peptides, peptoids, or carbohydrates in the practice of the instant invention.

-13-

The term "mixture" as used herein refers to a composition having a plurality of similar components in a single vessel.

5 The term "couple" as used herein refers to formation of a covalent bond.

The term "coupling moiety" refers to a soluble or insoluble support to which can be attached one or more active monomers and the corresponding encoding monomers. Insoluble supports ("solid support
10 means") may be any solid or semi-solid surface which is stable to the reaction conditions required for synthesis of the active and coding polymers, and is suitable for covalently attaching and immobilizing both polymers, for example, most resins commonly
15 employed in DNA and peptide synthesis, such as MBHA, Rink, and the like. The particular resin used will depend upon the choice of coding and active polymers and their associated synthetic chemistries. Soluble coupling moieties are molecules having functional
20 groups to which active and coding monomers may be attached. Each soluble coupling moiety must be able to accommodate at least one coding polymer and at least one active polymer, although the active and coding polymers need not be present in a 1:1 ratio.
25 The soluble coupling moiety may be as simple as an amino acid having an functional group in its side chain, or may be as complex as a functionalized (soluble) polymer.

The term "conjugate" as used herein refers
30 to the combination of any "active polymer" and its associated "coding" polymer. The conjugate may be formed using a "coupling moiety" or by binding both the "active polymer" and "encoding polymer" to the same support surface in close proximity with each
35 other so that the two polymers are "associated" with each other. When both polymers are bound to the same support surface, such as a small bead, the encoding polymer can be readily sequenced off of the bead and

-14-

the other "active polymers" remaining on the bead will be identified once the encoding sequence is known.

B. Related Libraries and Synthesis Methodologies for producing same.

There are many limitations with the current technologies for probing the receptor-binding properties of peptide libraries. Filamentous bacteriophage libraries offer the largest source of peptide diversity ($\approx 10^7$ - 10^8 different components) of any current technology to date (Scott, J. & Smith, G., Science, (1990), 249, 386-390; Devlin, J., Panganiban, L. & Devlin, P., Science, (1990), 249, 404-406; Cwirla, S., Peters, E., Barret, R. & Dower, W., Proc. Natl. Acad. Sci. U.S.A., (1990), 87, 6378-6382). These libraries, however, are limited to the natural set amino acids, suffer from biological biases (i.e., varying rates of growth, proteolysis, etc.) and also suffer, in practice, from high levels of background binding. The present invention is designed to overcome these difficulties.

Multiple-peptide synthesis technology has substantially increased the ability to generate individual peptides (Geysen, H., Meloen, R. & Barteling, S., Proc. Natl. Acad. Sci. U.S.A., (1984), 81, 3998-4002; Houghten, R., Proc. Natl. Acad. Sci. U.S.A., (1984), 5131-5135; Schnorrenberg, G. & Gerhardt, H., Tetrahedron, (1989), 45, 7759-7764; Gausepohl, H., Kraft, M., Boulin, C. & Frank, R. in Peptides: Chemistry, Structure and Biology (Proceedings of the 11th American Peptide Symposium, (1990), eds. Rivier J. & Marshall, G., (ESCOM, Leiden), pp. 1003-1004; Frank, R. & Döring, R., Tetrahedron, (1988), 44, 6031-6040; Fodor, S., Read, J., Pirrung, M., Stryer, L., Lu, A. & Solas, D., Science, (1991), 251, 767-773). The synthesis of $\approx 10^4$ individual peptides per cm^2 of glass wafer represents the diversity limit of this technology (Fodor, S.,

-15-

Read, J., Pirrung, M., Stryer, L., Lu, A. & Solas, D., Science, (1991), 251, 767-773). A mixed-resin algorithm method (Furka, A., Sebestyén, M., Asgedom M. & Dibó, G., Int. J. Peptide Protein Res., (1991), 37, 487-493) has recently been used to generate solution-phase libraries (Houghten, R., Pinilla, C., Blondelle, S., Appel, J., Dooley, C. & Cuervo, J., Nature, (1991), 354, 84-86) and resin-bound peptide libraries (Lam, K., Salmon, S., Hersh, E., Hruby, V., Kazmiersky, W. & Knapp, R., Nature, (1991), 354, 82-84) that contain $\approx 10^6$ and $\approx 10^7$ components, respectively. The solution-phase libraries offer the advantage of providing quantitative receptor-binding information (Zuckermann, R., Kerr, J., Siani, M., Banville, S. & Santi, D.V., Proc. Natl. Acad. Sci. U.S.A., 89, 4505-4509 (1992)). Furthermore, these libraries allow the affinity of the solution conformation of a ligand to be determined, a quantity that is essential for rational drug design. An apparatus for the automated synthesis of equimolar peptide mixtures is disclosed in Zuckermann, R.N., Kerr, J.M., Siani, M.A. & Banville, S.C., Int. J. Pep. Pro. Res., (1992), 40, 498-507.

The publications cited and discussed above can be used in producing the active or binding polymer which is used in producing the conjugate of the present invention. Accordingly, the disclosures of all of these publications are incorporated herein by reference in order to disclose peptide and peptoid synthesis methodology. Although the methodology discussed within these references is extremely valuable with respect to the production of large amounts of different types of binding polymers, the mixtures of polymers produced by this methodology are often so large and complex that there are many practical limitations with respect to their actual analysis and use. The present invention can be readily applied with such synthesis methodologies in

-16-

order to provide an efficient, commercially practical method of analyzing the proteins produced using such methodology.

Both the mixed-resin and solution-phase methods, however, do not allow incorporation of many non-standard amino acids because of the limitations of peptide analysis. Resin-bound peptide libraries, in particular, suffer from a relatively slow rate of analysis (peptide sequencing at 3 beads per day) and are limited in complexity to $\approx 10^7$ beads/ml. In order to generate a "complete" peptide library, there must be multiple copies (>10) of any given peptide sequence. This becomes problematic at the sequencing stage because the same "hit" sequence may appear multiple times. The alternative is to work with libraries that are not complete at the risk of losing sequences that bind.

When using the methodology of the present invention, the sequence of a biologically active protein can be determined even without isolating the protein of interest. This can be done by synthesizing large numbers of different proteins on large numbers of different support surfaces such as small beads. An encoding polymer is attached to beads to identify each protein. A sample to be tested is then brought into contact with the beads and the beads are observed with respect to which proteins bind to a receptor site in the sample. The bead having the receptor bound thereon is analyzed by sequencing the coding polymer which has also been synthesized on the bead. When the encoding polymer has been sequenced, the sequence of the active polymer, which may be a peptide, can be readily deduced. Thus, the present invention makes it possible to determine the activity and sequence an active polymer, such as a biologically active peptide, without ever isolating the peptide.

-17-

C. General Methodology

This invention describes a methodology for the synthesis and screening of large synthetic polymer libraries that contain non-standard amino acids and even non-amide based polymers. The strategy utilizes a modified mixed-resin peptide synthesis methodology to simultaneously synthesize two polymer sequences: one polymer strand (the "binding" strand) is synthesized for the intended purpose of receptor binding, and the second strand (the "coding" strand) contains standard amino acids or deoxyribonucleotides that encodes for the binding strand (Figure 1). The ability to decipher the binding sequence by analysis of the coding strand with standard peptide or oligonucleotide techniques allow the inclusion of a wide variety of novel building blocks and conformational constraints into a diverse ligand library.

This invention also describes a methodology to increase the size ($>10^8$) and screening rate of a ligand library. The method uses two polymers as above, but specifically utilizes an oligodeoxyribonucleotide for the "coding" strand. The use of DNA as the coding strand allows for an increased sensitivity of detection (fmol vs pmol for peptide analysis). This increased sensitivity allows for a larger library size since the amount of polymer needed for detection is reduced dramatically. The rate of sequence determination of receptor binders is increased since many samples can be analyzed in parallel.

In order to couple a polymer's sequence information with a peptide or oligonucleotide sequence, there needs to be a method that unambiguously correlates each polymer to each other. Thus, when any particular non-standard amino acid (or other monomer) is added to a "binding" polymer chain, the corresponding information (amino acid or

-18-

nucleotide monomer) must also be added to the "coding" strand. A "genetic code" is thus established (Table 1) where each binding monomer corresponds to (a multiple) of standard amino acids or nucleotides on the coding strand. For example, the use of three standard amino acids or nucleotides, in a 3:1 ratio with a novel monomer, would allow for the unambiguous representation of 27 novel monomers.

TABLE 1. Custom genetic code.

# of bases	"codon" length	# of monomers coded for
2	3	8
2	4	16
2	5	32
3	2	9
3	3	27
4	2	16
4	3	64

20

The synthesis of coded libraries requires a modified mixed-resin algorithm (Figure 2). The resin beads are divided into equal portions, a unique monomer is added to the "binding" strand, followed by the coupling of a corresponding amino acid or nucleotide to the "coding" strand. The resin aliquots are then combined to generate a mixture. A set of compatible protecting groups is thus required to preferentially deprotect and extend each strand independently.

Two synthesis formats are possible for amptide libraries, one that generates resin-bound libraries and one that generates solution-phase libraries (Figure 3). Resin-bound libraries can be synthesized using non-hydrolyzable linkers that are derivatized with the "binding" and "coding" monomers strands. Solution-phase libraries can be synthesized

-19-

as a 1:1 polymer:peptide/DNA conjugate via a hydrolyzable linker attached to the resin.

Peptide as the "Coding" strand

5 The use of base-labile Fmoc-protected monomers and acid-labile (N^{α} -Ddz-protected amino acids (Birrr, C., Nassal, M., Pipkorn, R., Int. J. Peptide Protein Res., (1979), 13, 287-295), for example, allow for selective deprotection and coupling to two
10 individual polymer strands. Resin-bound libraries can be generated by the derivatization of non-hydrolyzable resins with a 1:1 ratio (or any desired ratio) of Fmoc:Ddz monomers (Figure 4). This introduces two differently protected amino acids that can be extended
15 independently. Solution-phase libraries that contain a 1:1 ratio of binding:coding strands can be synthesized by using a hydrolyzable Fmoc-Lysine(Moz)-OH linker that allowed for chain growth at both the α - and ϵ -amino groups. Amino acids which do not contain
20 functional groups are preferred for the "coding" strand in order to minimize unwanted binding interactions.

 The receptor-binding ligand can be identified by bead staining techniques (Lam, K.,
25 Salmon, S., Herish, E., Hruby, V., Kazmiersky, W. & Knapp, R., Nature, (1991), 354, 82-84) and the sequence determined by N-terminal Edman degradation. In order to ensure that only the "coding" strand is sequenced, it is essential that the N-terminus of the
30 "binding" strand be acetylated or otherwise made non-sequencable.

DNA as the "Coding" Strand

35 The construction of libraries with DNA as the coding strand is similar to those with peptides but offers several advantages: the information storage and replicative properties of DNA allow for

-20-

increased sensitivity of detection, a larger library size and an increased rate of sequence determination.

The synthesis of DNA as the coding polymer requires compatibility between the assembly of Fmoc-based monomers and standard DNA chemistry. These synthesis strategies are likely to be compatible ((a) Juby, C., Richardson, C. & Brousseau, R., Tet. Letters, (1991), 32, 879-882. (b) Haralambidis, J., Duncan, L., Angus, B. & Tregear W., Nucleic Acid Res., (1990), 18, 493-499) (see Table 2). Alternatively, allyl-based protection strategies exists for both peptide (Lyttle, M.H.; Hudson, D., Peptides: Chemistry and Biology (Proceedings of the 12th American Peptide Symposium): Smith, J. and Rivier, J.E., Eds.; ESCOM, Leiden, 1992, pp. 583-584) and oligodeoxyribonucleotide (Hayakawa, Y., Wakabayashi, S., Kato, H. & Noyori, R., J. Am. Chem. Soc., (1990), 112, 1691-1696) synthesis. The assay of solution-phase libraries can be facilitated by using only pyrimidines in the coding strand, thereby avoiding the potential problem of base pairing between individual strands.

Strategies for the synthesis of coding and active polymers, as well as matching the active polymer with a coding sequence to provide a genetic tag are described in Brenner et al. Proc. Natl. Acad. Sci. U.S.A., 89:5381-5383 (June 1992) which is incorporated herein by reference.

30

35

-21-

TABLE 2

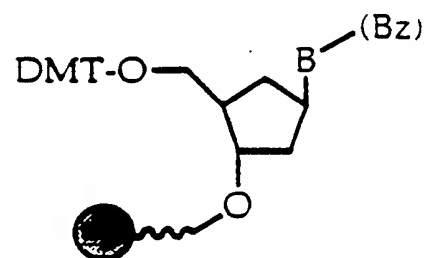
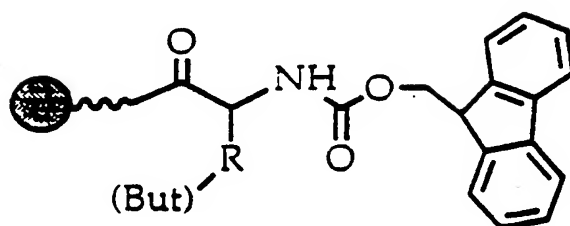
Compatibility of DNA vs. Peptide Synthesis Chemistry

5

Peptide chemistry

Oligonucleotide Chemistry

10



15

Permanent Protecting Groups:

-OH	t-butyl ether	A,C	benzoyl amide
-CO ₂ H	t-butyl ester	G	isobutyryl
-NH ₂	t-boc	-P=O	cyanoethyl, methyl
his, cys	trityl		
arg	sulfonyl		

20

removed by:

85% trifluoroacetic acid
2 hours @ room temp

conc. NH₄OH 55°C 5 hours

25

Temporary protecting groups:

9-fluorenylmethoxycarbonyl 4,4-dimethoxytrityl

removed by:

20% piperidine

3% trichloroacetic acid

30

Cleavage from solid support:

85% trifluoroacetic acid
2 hours @ room temp

conc. NH₄OH 55°C 5 hours

35

SUBSTITUTE SHEET

-22-

TABLE 2 (Cont)

Compatibility of DNA vs. Peptide Synthesis Chemistry

5	Std. conditions	Alternative	Problem addressed
	t-Butyl carbamate/ ester	allyl	amino acid deprotection from detritylation reagent
10	20% Piperidine	2% DBU	substitution at exocyclic amine
	cyanoethyl phosphate	allyl	acylation of phosphate
	NH ₄ OH	ethylenediamine/ EtOH	racemization
15	iodine	t-Butylhydro- peroxide	Tyr, Met, Cys oxidation
	controlled pore glass	polystyrene resin	resin transferring
	T, C & G	T, C, G & A	depurination
20	T, C	T, C, G & A	base pairing

25

30

35

-23-

Resin-bound libraries can be synthesized by using non-hydrolyzable linkers to attach both the C-terminus of the peptide and the 3'-end of the oligonucleotide to the same bead. Solution-phase

5 libraries can be synthesized as a 1:1 peptide-oligonucleotide conjugate, in which the C-terminus of the peptide is attached to the 3'-end of the oligonucleotide through a hydrolyzable Fmoc-Ser(O-Dmt) linker which is attached to the resin.

10 The identification of binders in the resin-bound peptide libraries can be detected by the bead staining methodology (Lam, K., Salmon, S., Hersh, E., Hruby, V., Kazmiersky, W. & Knapp, R., Nature, (1991), 354, 82-84). Although the peptides are bound to a

15 solid-phase, there does not have to be a 1:1 peptide-oligonucleotide ratio since the DNA can be amplified prior to the determination of its sequence. In fact, less DNA is preferred so that there will be less interference with the polymer's binding properties.

20 Once a bead is identified, the DNA sequence is determined (Stahl, S., Hultman, T., Olsson, A., Mois T., et al., Nucleic Acid Res., (1988), 16, 3025-3038) after PCR amplification or by thermal-cycle sequencing (Figure 8). This requires the inclusion of one or

25 more primer sites neighboring the coding region of the oligonucleotide. Similarly, the use of solution-phase libraries requires isolation of each sequence from each other. This can be accomplished by restricting the DNA after PCR amplification and inserting it into

30 M13 (or other suitable vector) for clonal isolation and sequencing (Figure 9).

EXAMPLES

The following examples will provide those

35 skilled in the art with a complete disclosure of how to make and use the invention and are not intended to limit the scope of the invention. Efforts have been made to insure accuracy with respect to numbers used

-24-

(e.g. amounts, temperature, etc.), but some experimental error and deviation should be accounted for. Unless indicated otherwise, parts or parts by weight, molecular weight is weight average molecular weight, temperature is in degrees centigrade and pressure is at or near atmospheric.

Example 1

The independent synthesis of two unambiguously correlated sequences has been successfully completed. The subsequent sequence analysis of the "coding" strand has also been demonstrated. For convenience, two peptide sequences were chosen. The "binding" strand was synthesized with N^α-Fmoc-protected amino acids and the "coding" strand was synthesized with N^α-Ddz-protected amino acids. The simultaneous synthesis of the two peptide strands was tested in two formats, 1) resin-bound peptide library synthesis and 2) solution-phase peptide libraries using a hydrolyzable Fmoc-Lys(Moz)-OH linker (Wang, S.S.; Chen, S.T, Wang, K.T., and Merrifield, R.B., Int. J. Peptide Protein Res., (1987), 30, 662-667). These syntheses were performed on single peptides (not libraries) as a demonstration of research concept and in order to allow the full characterization of the synthesis products.

30

35

-25-

A. Synthesis of a Resin-Bound Library Model

"Binding" Sequence: Ac-Arg-Leu-Val-Thr-His
(Fmoc peptide)

Resin

5

"Coding" Sequence: H₂N-Ala-Ser-Gly-Glu-Phe-Ala
(Ddz peptide)

Synthesis Scheme:

Step #	Description
1	Derivatization of MBHA resin with 1:1-Ddz-Ala-OH: Fmoc-His(Trt)-OH
2	Deprotection of "binding" strand with 20% piperidine/DMF
3	Coupling of second "binding" amino acid, Fmoc-Thr(But)-OH
4	Deprotection of "coding strand" with 7.5% TFA/CH ₂ Cl ₂
5	Coupling of second "coding" amino acid, Ddz-Phe-OH
6	Repeat steps 2-5 with alternative Fmoc and Ddz deprotection and corresponding amino acid coupling
7	Final Deprotection of Fmoc followed by acetylation
8	TFA deprotection of side-chain groups

25

After the TFA deprotection, the model library bead has two independently synthesized sequences and is ready for assay. Only the coding strand has a free α -amino group and can be characterized by N-terminal Edman degradation. The binding strand is acetylated and therefore will not interfere with the sequencing. The two peptides were cleaved from the resin with HF thereby providing both the "binding" and "coding" sequences as free peptides. The amino acid composition, mass spectroscopy and N-terminal sequencing data are consistent with the correct products. (See Figures 5, 6 and 7.)

30

35

-26-

Mass Spectrometry:

		theoretical	observed
	Ddz peptide	580.2	580.2
5	Fmoc peptide	666.9	666.4

Amino Acid Composition:

	Fmoc Peptide			Ddz Peptide		
10		Theor. (%)	Observ. (%)		Theor. (%)	Observ. (%)
	His	20	21	Ala	33	33
	Thr	20	20	Glu	17	17
	Val	20	19	Phe	17	17
15	Leu	20	29	Gly	17	15
	Arg	20	20	Ser	17	15

N-Terminal Edman Sequencing of Resin beads (coding peptide only):

20	Cycle #	amino acid	pmol
	1	Ala	45
	2	Ser	12
	3	Gly	30
25	4	Glu	28
	5	Phe	28
	6	Ala	29

Example 2

30

Synthesis of a Solution-Phase Peptide Library Model

In this example, a 1:1 solution-phase adduct between a "binding" and a "coding" strand was synthesized and fully characterized. The "binding" strand was assembled with Fmoc-protected monomers, and the "coding" strand was assembled with Ddz-protected monomers.

-27-

"Binding" sequence: Ac-Glu-Ser-Thr-Arg-Pro-nLeu-Lys-B-
(Fmoc peptide) Ala-NH₂

"Coding" sequence: H₂N-Gly-Ala-Phe-Gly-Ala-Phe-CONH
(Ddz peptide)

5

Synthesis Scheme:

Step #	Description
1	Derivatization of Rink Resin with Fmoc-B-Ala-OH spacer
2	Fmoc deprotection with 20% piperidine/DMF
3	Coupling of Fmoc-Lys (4-methoxybenzyloxycarbonyl)-OH
4	Fmoc deprotection of "binding" strand with 20% piperidine/DMF
5	Coupling of first "binding" monomer Fmoc-nLeu-OH
6	Cleavage of "coding" strand 4-Methoxybenzyloxycarbonyl (Moz) group with 7.5% TFA/CH ₂ Cl ₂
7	Coupling of first "coding" amino acid Ddz-Phe-OH
8	Repeat steps 4-7 with corresponding amino acids
9	Final Fmoc deprotection and acetylation
10	Cleavage and deprotection of resin sample with TFA

Following TFA cleavage and deprotection, the model solution-phase library contains a 1:1 Fmoc/Ddz conjugate peptide. One peptide sequence was synthesized and not a mixture in order to fully characterize the reaction product. The amino acid composition and mass spectroscopy data are consistent with the correct product. In addition, the "binding" and "coding" hybrid peptides were tested in a competition ELISA format. The ELSTRPnL "binding" sequence binds to an anti-gp120 antibody with submicromolar affinity. This value was not affected by the presence of the "coding" peptide.

-28-

Mass Spectroscopy:

	Theoretical	Observed
Fmoc/Ddz peptide conjugate:	1492.7	1492.6

5 Amino Acid Composition:

Fmoc Peptide			Ddz Peptide		
	Theor. (%)	Observ. (%)		Theor. (%)	Observ. (%)
10 Glu	7.1	7.0	Phe	4.2	14.7
Ser	7.1	7.1	Ala	14.2	14.4
Thr	7.1	7.2	Gly	14.2	13.8
Arg	7.1	7.2			
Pro	7.1	7.9			
15 Nleu	7.1	6.3			

 The instant invention is shown and described
herein in what is considered to be the most practical,
and preferred embodiments. It is recognized, however,
20 that departures may be made therefrom which are within
the scope of the invention, and that obvious
modifications will occur to one skilled in the art
upon reading this disclosure.

25

30

35

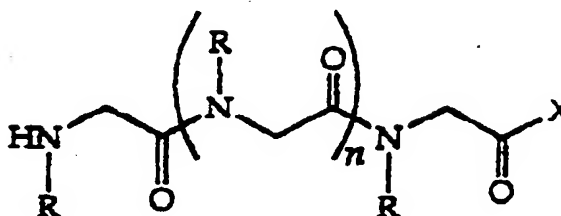
-29-

CLAIMS

1. An assay conjugate, comprising:
an active polymer comprising monomers
5 selected from the group consisting of peptide and
peptoid monomers;
an encoding polymer comprising encoding
monomers, wherein the encoding polymer corresponds to
and allows identification of the active polymer; and
10 a coupling moiety covalently coupled to the
active peptide and the encoding polymer.
2. The conjugate of claim 1, wherein the
coupling moiety comprises a solid support.
15
3. The conjugate of claim 1, wherein the
coupling moiety comprises a soluble linking group.
4. The conjugate of claim 1, wherein all
20 active polymers coupled to a single selected solid
support are identical.
5. The conjugate of claim 1, wherein the
active polymer comprises a polypeptide and the
25 encoding polymer comprises a DNA or RNA oligonucleo-
tide.
6. The conjugate of claim 1, wherein the
active polymer comprises a polypeptoid and the
30 encoding polymer comprises a DNA or RNA oligonucleo-
tide.

-30-

7. The conjugate of claim 6, wherein the polypeptoid comprises a polymer of monomers of the formula:



wherein R is alkyl of 2-6 carbon atoms, haloalkyl of 1-6 carbon atoms wherein halo is F, Cl, Br, or I, alkenyl of 2-6 carbon atoms, alkynyl of 2-6 carbon atoms, cycloalkyl of 3-8 carbon atoms, alkoxyalkyl of 2-8 carbon atoms, aryl of 6-10 carbon atoms, arylalkyl of 7-12 carbon atoms, arylalkyl of 7-12 carbon atoms substituted with 1-3 radicals independently selected from halo and nitro and hydroxy, aminoalkyl of 1-6 carbon atoms, hydroxyalkyl of 1-6 carbon atoms, carboxy, carboxyalkyl of 2-6 carbon atoms, carboalkoxy-alkyl of 3-10 carbon atoms, carbamyl, carbamylalkyl of 2-6 carbon atoms, imidazolyl, imidazolylalkyl of 4-10 carbon atoms, pyridyl, pyridylalkyl of 6-10 carbon atoms, piperidyl, piperidylalkyl of 5-10 carbon atoms, indolyl, or indolylalkyl of 9-15 carbon atoms.

8. A mixture of conjugates of claim 1, wherein the mixture comprises at least two distinct active polymers, the coupling moiety is a solid support, distinct active polymers are covalently coupled to separate solid supports and a distinct encoding polymer corresponding to each active polymer is covalently coupled to the support coupled to its corresponding active polymer.

-31-

9. The mixture of claim 8, wherein the mixture comprises at least ten distinct active polymers.

5 10. The mixture of claim 9, wherein the mixture comprises at least 100 distinct active polymers.

10 11. A conjugate, comprising:
a biologically active peptide comprised of five or more amino acids;
an encoding polymer comprised of nucleic acids wherein one or more nucleic acids within the encoding polymer correspond to and can be readily
15 identified with the amino acids of the active polymer;
and
a coupling moiety covalently coupled to the active peptide and the encoding polymer.

20 12. The conjugate of claim 11, wherein the amino acids are selected from the group consisting of alanine, cysteine, aspartic acid, glutamic acid, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, proline, glutamine,
25 arginine, serine, threonine, valine, tryptophan, and tyrosine.

30 13. The conjugate of claim 11, wherein the encoding polymer includes three nucleotides for each amino acid of the active polymer and the three nucleotides are nucleotides which naturally encode the corresponding amino acid of the active polymer.

35 14. The conjugate of claim 11, wherein the coupling moiety comprises a solid support.

-32-

15. The conjugate of claim 14, wherein the solid support is in the form of a spherical bead having a diameter of less than one centimeter.

5 16. The conjugate of claim 11, wherein the coupling moiety comprises a soluble linking group.

10 17. A method of synthesizing an encoded peptide or peptoid polymer, the method comprising:
a) providing a coupling moiety;
b) coupling to the coupling moiety an active monomer, and coupling to the coupling moiety an encoding monomer which corresponds to the active monomer
15 to form a conjugate having a bound active monomer and a bound encoding monomer; and
c) repeating step b) until an active polymer of the desired length is
20 obtained.

18. The method of claim 17 wherein the active monomers are amino acids and the encoding monomers are nucleic acid bases.
25

19. A method of synthesizing a mixture of encoded polymers, the method comprising:
a) providing a plurality of coupling moieties;
30 b) dividing the plurality of coupling moieties into a plurality of aliquots;
c) within each aliquot, coupling to each coupling moiety an active monomer, and coupling to each said coupling moiety an encoding monomer which corresponds
35 to the active monomer to form a plurality of conjugates having a bound active monomer and a bound encoding

-33-

monomer, each in a separate aliquot,
wherein different aliquots may contain
different active monomers;

- 5 d) combining the aliquots of conjugates to
 form a mixture of conjugates; and
 e) repeating steps b-d) until a mixture of
 active polymers of the desired length
 is obtained.

10 20. The method of claim 19, wherein the
 active monomers comprise activated amino acids and the
 coupling moiety comprises a solid support.

15

20

25

30

35

1 / 9 -

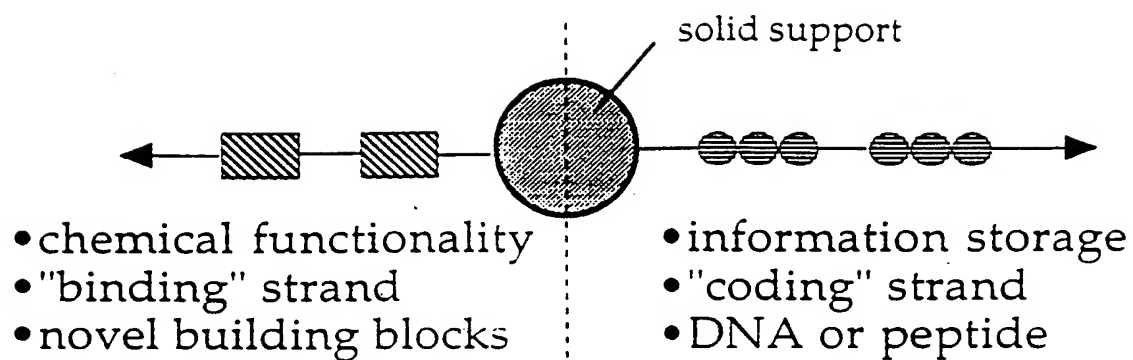


FIGURE 1

- 2 / 9 -

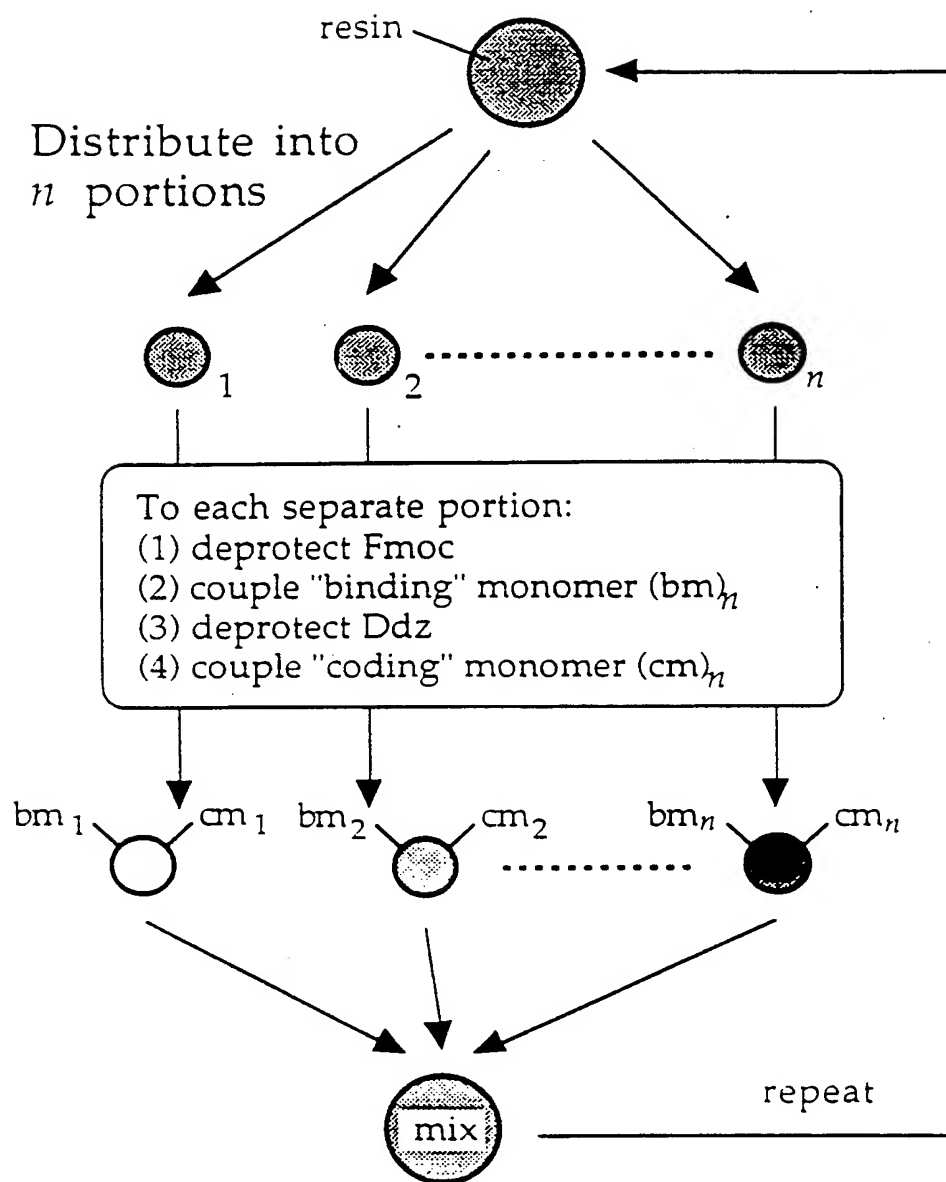


FIGURE 2

- 3 / 9 -

(1) Resin-bound:

polymer/DNA

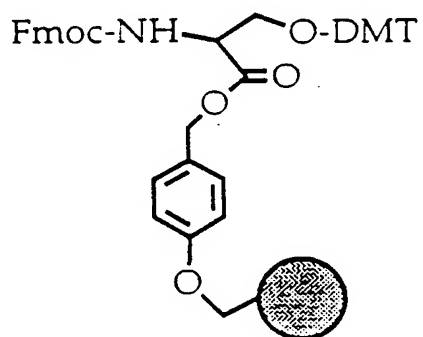


polymer/peptide



(2) Solution-phase conjugate:

polymer/DNA



polymer/peptide

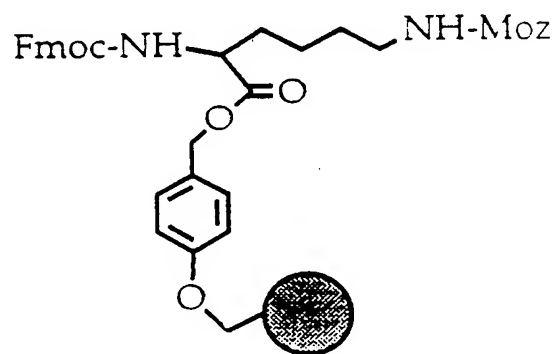


FIGURE 3

4/9

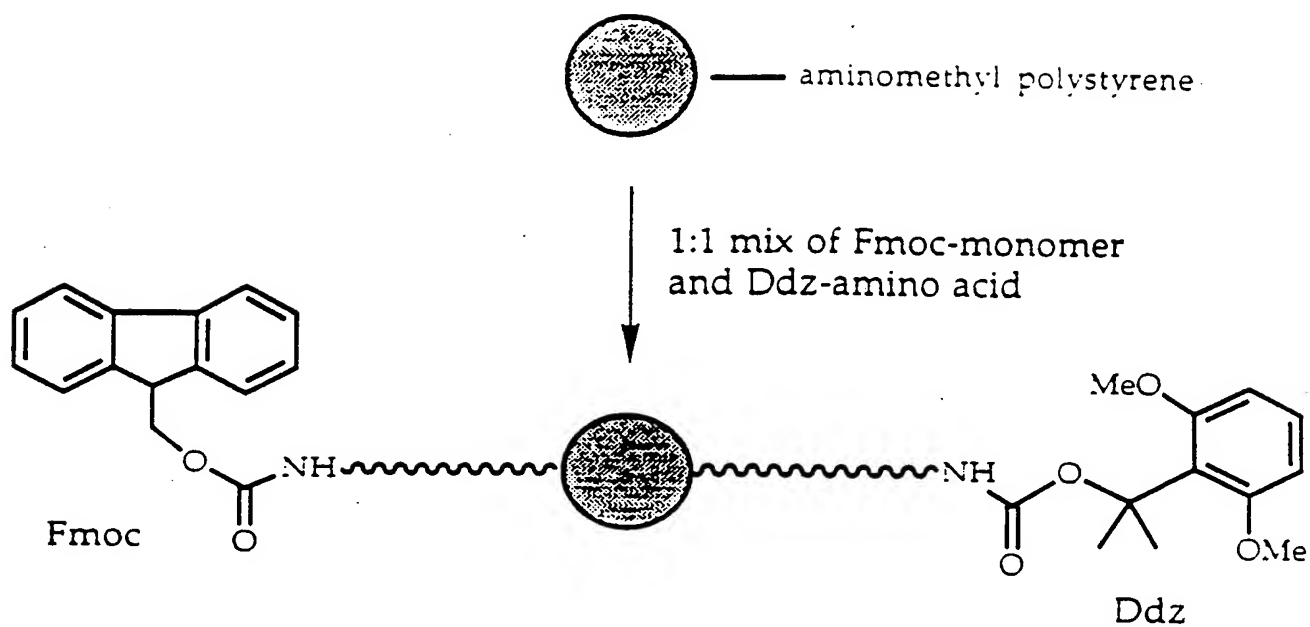


FIGURE 4

- 5 / 9 -

HPLC chromatogram of HF cleaved/deprotected resin sample:

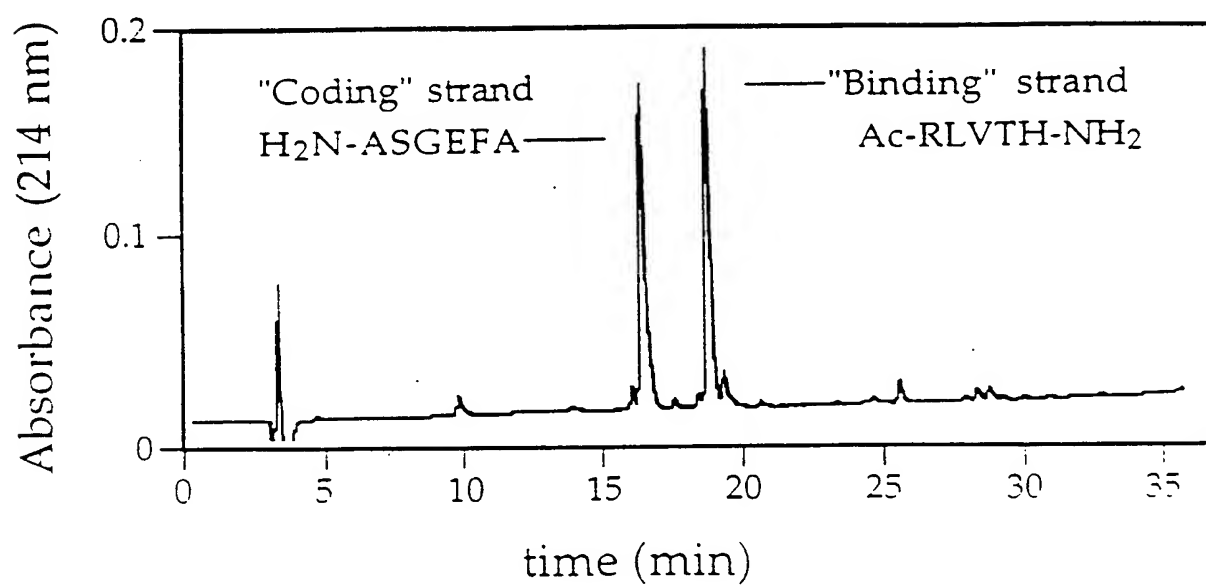


FIGURE 5

- 6 / 9 -

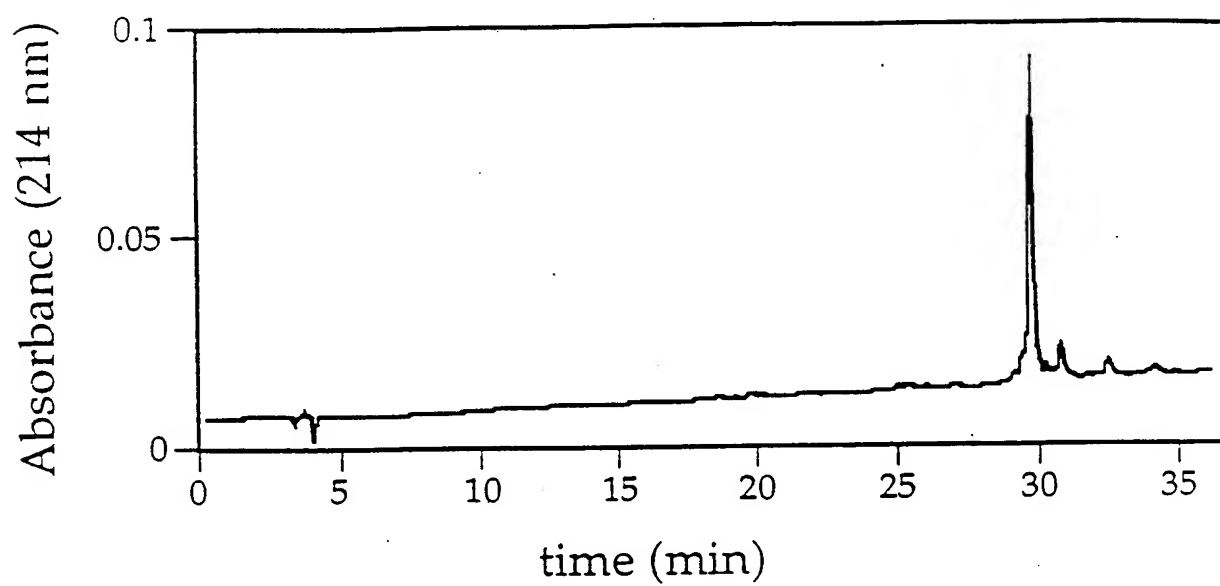
HPLC chromatogram of TFA cleaved/deprotected peptide:

FIGURE 6

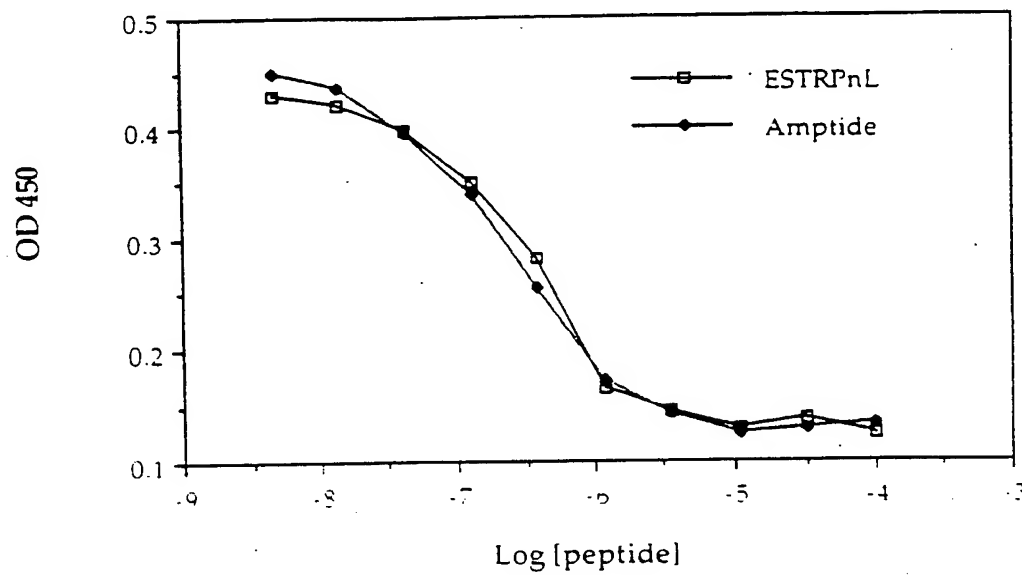
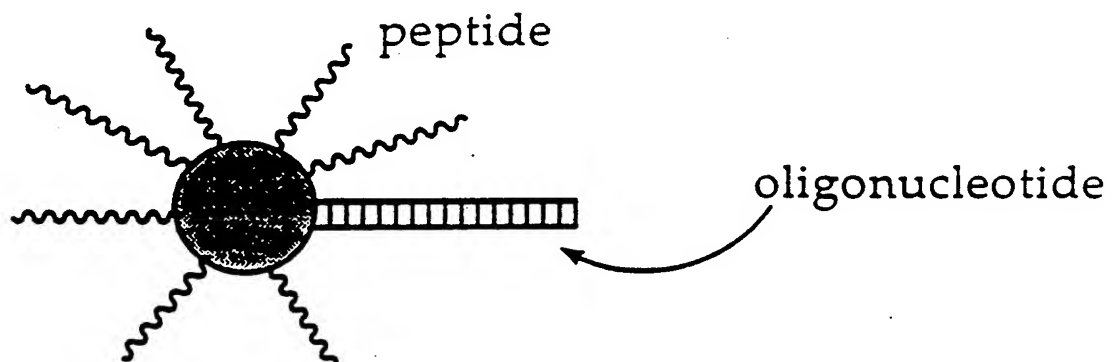
Competition ELISA of "binding" sequence vs. "binding/coding" sequence

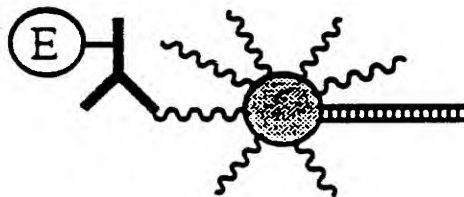
FIGURE 7

Analysis of solid-phase Amptides

- simultaneous synthesis of peptide & corresponding oligo



- detection/isolation of bead of interest



- DNA sequencing/amplification directly on a solid-phase*

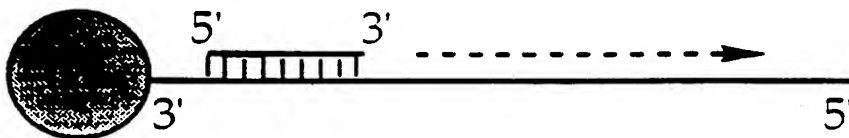


FIGURE 8

9/9

Analysis of solution-phase amptides

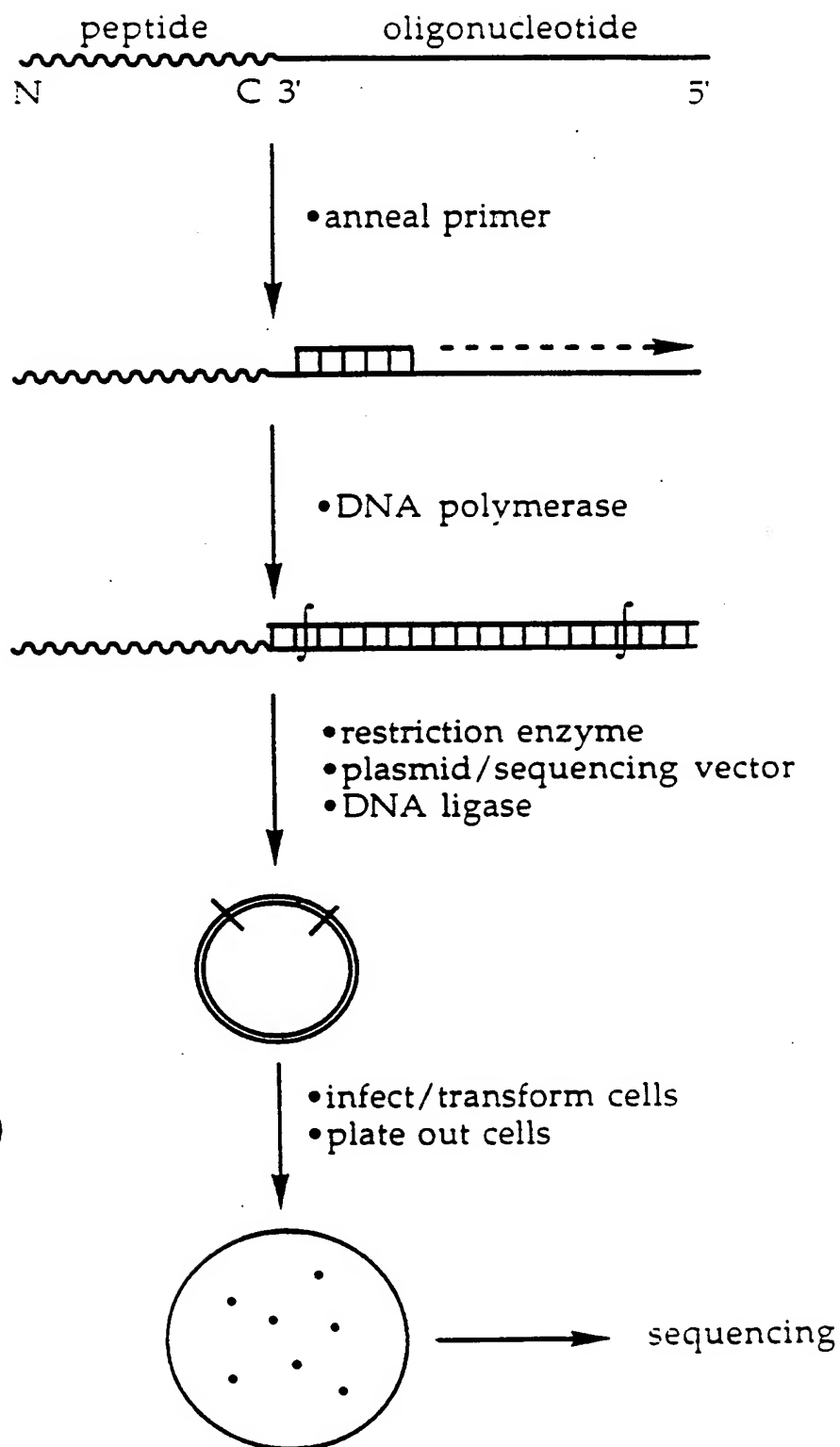


FIGURE 9

INTERNATIONAL SEARCH REPORT

 Int. national application No.
PCT/US93/12013
A. CLASSIFICATION OF SUBJECT MATTER
 IPC(5) : C07C 237/00; C07H 21/00; C07K 1/00, 5/00, 7/00, 13/00, 17/00
 US CL : 530/300, 350, 333, 334, 338, 402; 536/25.3; 564/159

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/300, 350, 333, 334, 338, 402; 536/25.3; 564/159

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

search terms: oligonucleotide, DNA, peptide, synthesis, combinatorial

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Proceedings of the National Academy of Sciences U.S.A., Vol. 89, issued June 1992, S. Brenner et al., "Encoded combinatorial chemistry", pages 5381-5383, see the entire document.	1, 2, 4, 5, 8-12, 14, 15, 17-20 ----- 3, 6, 7, 13, 16
X --- Y	Science, Vol. 257, issued 17 July 1992, I. Amato, "Speeding up a chemical game of chance", pages 330-331, see the entire document.	1, 2, 4, 5, 8-12, 14, 15, 17-20 ----- 3, 6, 7, 13, 16

☒ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"A" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 01 February 1994	Date of mailing of the international search report 24 FEB 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer ERIC GRIMES <i>D. Kuyza for</i> Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US93/12013

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P ---- Y, P	Proceedings of the National Academy of Sciences U.S.A., Vol. 90, issued November 1993, M. C. Needels et al., "Generation and screening of an oligonucleotide-encoded synthetic peptide library", pages 10700-10704, see the entire document.	1, 2, 4, 5, 8-12, 14, 15, 17-20 ----- 3, 6, 7, 13, 16
Y	Science, Vol. 254, issued 06 December 1991, P. E. Nielsen et al., "Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide", pages 1497-1500, see page 1498.	6, 7
Y	BioTechniques, Vol. 13, No. 3, issued September 1992, R. A. Houghten et al., "The use of synthetic peptide combinatorial libraries for the identification of bioactive peptides", pages 412-421, see the entire document.	1-20
Y	US, A, 4,359,353 (KYDD) 16 November 1982, see columns 2-3.	1-20
A	Nature, Vol. 354, issued 07 November 1991, K. S. Lam et al., "A new type of synthetic peptide library for identifying ligand-binding activity", pages 82-84.	1-20
A	Science, Vol. 249, issued 27 July 1990, J. J. Devlin et al., "Random peptide libraries: a source of specific protein binding molecules", pages 404-406.	1-20